ED 334 260 TM 016 870

AUTHOR Ingels, Steven J.; Scott, Leslie A.
TITLE National Dropout Statistics from a Longitudinal Cohort Perspective: Estimating Rates of School-Leaving and School Noncompletion.
INSTITUTION National Opinion Research Center, Chicago, Ill.
SPONS AGENCY National Center for Education Statistics (ED), Washington, DC.
PUB DATE Apr 91
CONTRACT NCES-RS-88-3001
NOTE 22p.; Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, IL, April 3-7, 1991).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Cohort Analysis; Data Collection; Demography; *Dropouts; Enrollment; *Estimation (Mathematics); Grade 8; Longitudinal Studies; *National Surveys; *Predictive Measurement; Secondary Education; *Statistical Data; Trend Analysis
IDENTIFIERS High School and Beyond (NCES); *National Education Longitudinal Study 1988

ABSTRACT

The collection of dropout statistics is discussed from the perspective of the National Education Longitudinal Study of 1988 (NELS:88). The NELS:88 follows a cohort of 1988 eighth graders over time and is designed to provide trend data about transitions experienced as students progress through the educational system. Dropout statistics are provided for each year, for those who have not finished school at any given point and for the cohort as a whole. The concept of a dropout has been operationalized to distinguish those who are not enrolled in some alternative education. The ways in which the NELS:88 remedies weaknesses of the High School and Beyond study are reviewed, centering on the more accurate estimates provided by confirmation from school enrollment and demographic data. Data from longitudinal cohort studies similar to the NELS:88 make more meaningful analysis of educational trends possible. Three appendices provide supplemental information about survey methodology. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED334260

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

STEVEN J. INGELS

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

NATIONAL DROPOUT STATISTICS FROM A LONGITUDINAL COHORT PERSPECTIVE: Estimating Rates of School-Leaving and School Noncompletion

Steven J. Ingels and Leslie A. Scott

National Opinion Research Center (NORC)
University of Chicago
1155 East 60th Street
Chicago, Illinois 60637-2799

SYMPOSIUM ON DROPOUT STATISTICS FROM ADMINISTRATIVE RECORDS
American Educational Research Association Annual Meeting
Chicago, Illinois, April 6, 1991, Session 7.04

This paper is based on the National Education Longitudinal Study of 1988, sponsored by the National Center for Education Statistics (NCES), Office of Educational Research and Improvement, U.S. Department of Education. Dr. Jeffrey A. Owings is the Project Officer for the NELS:88 First Follow-Up. The ideas and opinions expressed in this paper are those of the authors and do not necessarily represent positions or conclusions endorsed by NCES.

NORC Project 4492; NCES Contract RS-8800-3001

TM016870

BEST COPY AVAILABLE

1. Introduction: Issues and Background

1.1 Issues.

While papers in the symposium generally illuminate the twin strands of capacities of administrative records systems per se, and differences between state and local records systems and implications of those differences for achieving uniform statistical reporting, the primary thrust of this presentation complements rather than recapitulates these two concerns. In general, household surveys such as the Current Population Survey and school-based longitudinal studies such as NELS:88 offer a check on and corrective to estimates produced from administrative records sources. Such studies also, however, confront the same definitional issues as do records systems. Both as a supplementary source, and as alternative strategies for obtaining valid and reliable dropout data, cross-sectional household and student longitudinal surveys have much to contribute to methodologies for the pursuit of accuracy in administrative records. In addition, the complementarity of records-based approaches, and longitudinal and household survey approaches, will be increased to the extent that taxonomies of key phenomena respect a common definitional language that ensures the potential for comparability between measures based on divergent methodologies and diverse designs.

The National Education Longitudinal Study of 1988 (NELS:88) follows a panel of spring 1988 eighth graders over time. All dropouts identified in the study are followed with certainty, and complete a special dropout questionnaire and cognitive tests. The study gathers data on enrollment status and high school and equivalency completion, as well as school records-based reports on enrollment status and (through transcripts) school completion. NELS:88 is designed to produce national estimates--overall, and for policy relevant subgroups (blacks, Hispanics, Asians)--and estimates for Census regions and divisions.

NELS:88 also permits cross-checking of the school's records-based enrollment status reports with an external source (the household of the putative dropout). Thus, while the longitudinal approach described in this paper places considerable reliance on administrative records, it also seeks corroboration of records from non-records sources, and amasses data at several levels, and over various time points, that can be combined and recombined to match a variety of definitions and statistical reporting and policy analysis purposes.

After laying out initial background information, this paper addresses three kinds of issues. First, it talks about issues of dropout definition and the calculation of dropout rates in relation to NELS:88--what kind of dropout statistics, in other words, does a longitudinal study produce?

Second, it examines methodological issues. Specifically--how are dropout definitions to be operationalized? How is dropout status to be confirmed?

Third, it considers gaps and weaknesses in the dropout statistics collected by the High School and Beyond (HS&B) study, and shows how NELS:88 is designed to

overcome these weaknesses. Specifically, it considers the late beginning point of HS&B (second semester of tenth grade), the problem of ineligibility exclusions, and the problems for estimation posed by sample nonparticipants, and shows how these problems will be addressed in NELS:88.

The overall point to be made about the NELS:88 approach to dropout statistics is the need, at this stage in our attempt to gather more comprehensive data about a complex phenomenon, for multiple yet comparable approaches to statistical reporting. These approaches need to take cognizance of the movement toward uniform definitions and categorization criteria, and gather and report data in a manner that maximizes comparability with more uniform definitions. The close comparison of the results of differing approaches is extremely important to the process of refining estimates and correcting for the limitations of each approach. At the same time that comparability with new, more uniform definitional criteria must be sought, data must also be captured and categorized in a way that maintains continuity and comparability with major historical sources--a requirement that is especially compelling for a study such as NELS:88, which is designed to provide trend data and support cross-cohort comparisons.

1.2 Historical Background: NCES's National Education Longitudinal Studies

The NELS program currently comprises three major studies: the National Longitudinal Study of the High School Class of 1972 (NLS-72); High School and Beyond (HS&B); and the National Education Longitudinal Study of 1988 (NELS:88). Taken together, these studies represent the educational experience of youth from three decades--the 1970s, 1980s, and 1990s. While NLS-72 sampled only high school seniors, two of these studies--HS&B, in its sophomore cohort--and NELS:88, with its eighth grade starting point--have a particular contribution to make to the collection and refinement of national dropout statistics, as well as to analyses of the dynamics and consequences of dropping out.

HS&B base year data collection was conducted in the spring of 1980. Students were selected using a two-stage probability sample with schools as the first-stage units and students within schools as the second-stage units. There were 1,015 public, private, and church-affiliated secondary schools in the sample and a total of 58,270 participating students. Unlike NLS-72, HS&B included cohorts of both tenth graders and twelfth graders. Since the base year data collection in 1980, three follow-ups of the HS&B cohorts have been completed, one in the spring of 1982, one in the spring of 1984, and the last in the spring of 1986. A fourth follow-up, of the sophomore cohort only, is planned for 1992.

The National Education Longitudinal Study of 1988 (NELS:88) is designed to provide trend data about critical transitions experienced by students as they leave elementary school and progress through high school and into college or their careers. Policy-relevant data about educational processes and outcomes will be collected over time, especially as it pertains to student learning, early and late predictors of dropping out, and school effects on students' access to programs and equal opportunity to learn. Data was collected from students, parents, teachers and school administrators in the spring of 1988, following a two-stage (schools, and then students) selection of the sample, that followed the basic model of HS&B, but with more extensive oversampling of private schools, as

well as of Asians and Hispanics. Baseline analyses of factors thought to be predictive of dropping out prior to school completion are reported in Hafner et al. (1990).

A subsample of approximately 20,000 of the 1988 eighth grade cohort of over 26,000 students is being followed at two-year intervals as this group passes through high school. Selected sample members will be further followed in 1994 and thereafter, into postsecondary education and the labor force. (All dropouts will be followed, however; no subsampling will take place with school leavers.) Data collection for the First Follow-Up was conducted in the spring of 1990, including completion of questionnaires and a cognitive test battery by both in-school sample members and those who had dropped out between 1988 and 1990. The study will support multivariate event-history analysis of dropout behaviors, and in particular, ascertain how many dropouts eventually complete school or obtain equivalent qualifications.

The field test for the NELS:88 Second Follow-Up is currently in progress, and tracing for the Second Follow-Up main study (data collection is scheduled for the spring of 1992) is also underway. The Second Follow-Up includes surveys of students, their parents, their teachers and school principals, as well as extensive records data collection, primarily in the form of secondary school transcripts, course offerings information, and course enrollment figures.

2. Definitional Issues

2.1 Dropout Definitions and the Calculation of Dropout Rates.

Event, Status and Cohort Definitions of NELS:88 Dropouts.

Event Definition: Tallying New Dropout Events. The event dropout rate is the proportion of students enrolled one year ago who have since dropped out of school (Frase, 1989). The Current Population Survey employs this measure, calculating the proportion of dropouts during the twelve-month period from October of one year to October of the next (Kominski, 1990), and is thus a rich (indeed, the sole) source of annual time-series data.

In NELS:88, the starting point is enrollment status in the spring of 1988. Eighth graders in 1988 were followed at three distinct stages in the NELS:88 First Follow-Up; a similar scheme will be employed in the Second Follow-Up.

Phase 1: Tracing; spring 1989 (99% of eighth grade cohort members successfully traced and enrollment status ascertained).
Phase 2: Autumn school contacting; fall 1989 (verifying school enrollment, freshening the sample).
Phase 3: Data collection; spring 1990 (reverification of school enrollment status).
Phase 4: Cleanup--post-data collection screening of ambiguous pending cases (December 1990-March 1991, but reference period = spring of 1990; thus phase 4, while operationally distinct, is temporally identical in its reference to phase 3)

However, during each of the three temporally distinct phases during which enrollment status data are gathered, an event date (for dropping out, and for returning to school) is also pursued. The three phases of the study, then, are essentially discovery points; a dropout event learned about in the spring of 1989 could have occurred at the end of the spring term in 1988. Since dropout dates will be collected (as well as drop in [return to school] dates for stopouts, that is, temporary dropouts), analysts will enjoy considerable flexibility in defining reference periods for event history analyses.

One could employ NELS:88 data to examine the proportion of students who dropped out between spring 1988 and spring 1989, and between spring 1989 and spring 1990. However, these anchor points are rather inexact. While the Census Current Population Survey (CPS), for example, works within a single month, HS&B and NELS:88 employ a broader temporal reference of about four months. Thus, we know the status of sample members as of the spring of 1989, the autumn of 1989, and the spring of 1990. Estimates derived from these broad reference periods may not correspond to estimates generated from a narrower reference period--say one month, or one particular day.[1] A sample member who drops out in March will be captured as a dropout if the school's Survey Day is held in May but not if the survey session occurs in February. In the latter case, one would, however, to obtain the dropout event date in the next (follow-up) round of the study.

While NELS:88 employs an event history definition of dropping out, for a number of reasons dropout spells of short duration may be missed by this methodology. For example, a student who was at an eighth grade survey day in February of 1988 may have been a dropout between March and June of 1988. If that student has returned to school in the autumn of 1988, the tracing phase would identify that student as in school, with a resultant underreporting of dropout/stopout events. Nevertheless, ascertaining status at three points in the two-year period between follow-ups serves to maximize the number of dropout events recorded. Moreover, NELS:88--like HS&B before it--offers the opportunity to inquire on the student questionnaire whether there have been gaps in school attendance of a given duration (such as four consecutive weeks of unexcused absence). The October to October definitions employed by CPS and proposed for CCD cannot capture within-school-year stopout phenomena.

Status Dropout Rate: Cumulative Measures of Current School Non-Completers. The status dropout rate represents the number of dropouts who have not finished high school at any given time, regardless of when they left school. The status rate can, of course, be applied to age subsets of the population (that is, age cohorts) or to the population as a whole. It can also be applied to a grade cohort, when anchored to a specific time point (for example, as in the case of

---

[1]There may also be, as Hafner points out (1990, p.6) a more general comparison problem between various studies that collect dropout data at differing points in the year. She notes that October enrollments tend to be inflated, spring enrollments lower. Nevertheless, it is a separate question, for which there is at present little empirical evidence to suggest an answer, whether spring to spring or autumn to autumn measurements produce significantly different dropout <u>rates</u>.

NELS:88, which deals with 1987--88 eighth graders in following years without regard to whether they remain in modal grade sequence or not).

Cohort Dropout Rate. The cohort rate measures what happens to a single defined group of students over a period of time. It thus offers an opportunity to both measure the cumulative percentage of individuals in the cohort who ever drop out, as well as to ascertain the proportion of those who ever (if the study continues for long enough) return to school and who finally complete their schooling or obtain equivalency certificates. With cohort data, one can readily move back and forth, as needed, between what Barro and Kolstad (1987, p.9) call gross (any student who has ever dropped out) and net (students who are dropouts at a particular time) definitions. One can accommodate the fact as well that a given individual may drop out (and return to school) multiple times.

While NELS:88 readily supports calculation of a cohort dropout rate, it is important to note that in fact three cohorts are derivable from the study. While the HS&B sophomore cohort is a representative sample of 1980 high school sophomores who were resurveyed in 1982 and thereafter, the cohort is not representative of 1982 high school seniors. In contrast, NELS:88 underwrites three longitudinal cohorts: a representative sample of spring 1988 eighth graders; a representative sample of spring 1990 high school sophomores; and a representative sample of spring 1992 high school seniors. Each panel is augmented in a process of sample "freshening" (for an explanation of sample freshening, see Appendix A), which gives 1990 tenth graders and 1992 twelfth graders who were not in the eighth grade in the United States in the 1987-88 school year some chance of selection into the follow-up samples.

For purposes of calculation of dropout statistics a representative twelfth grade cohort might at first seem superfluous. Nonetheless, there will be 1992 twelfth graders who drop out in the course of the school year, some of them after their school's survey day. For the eighth grade cohort, most sample members will be in tenth grade in spring 1990, but some may still be in eighth grade, others in ninth or eleventh grade. By 1992, members of the eighth grade cohort may be in eighth grade or any grade higher, or may be early graduates; some will have returned to school after spells of dropping out. The tenth grade cohort will, in the same way, embrace those who remain in modal sequence, and those whose educational histories depart from it.

Perhaps the primary problem to be faced in NELS:88 in calculating a cohort dropout rate is just how inclusive of the full population that cohort rate should be. Later we will discuss the issue of whether the eighth grade cohort should encompass the full population of eligible participants and nonparticipants and ineligible students as well, or whether the cohort should be construed in more limited terms (as was the case in HS&B).

3. Methodological Issues

3.1 Operationalizing Dropout Definitions in HS&B and NELS:88

3.1.1 HS&B Operational Definition of a Dropout.

HS&B Base Year. In the HS&B base year, the student sample frame was derived from the eligible school sample. (As in NELS:88, certain types of special schools were excluded.) Within the eligible schools, a sophomore was defined as a student who expected to complete his/her tenth grade course work between April 1, 1980 and August 31, 1980. This was to include those students who might be held back or who might repeat tenth grade, but to exclude students dropping out before administration of the HS&B questionnaire in the spring of 1980. Thus any student who dropped out prior to the school's survey day in the spring of 1980 (survey sessions were held between February and June) was considered ineligible for the HS&B sample and was excluded. For purposes of ineligibility, the following definition of dropping out was used in the HS&B base year: a student who has been out of school for 20 or more consecutive days, is at least sixteen years of age, and is not expected to return, should be considered ineligible by virtue of being a dropout prior to survey day. Students with 20 are more consecutive absences who were not expected to return but were not yet sixteen years old were also excluded, though they were categorized as "lost students" rather than dropouts, since generally they had not reached the legal school-leaving age.

HS&B First Follow-Up. The dropout definition employed in the HS&B First Follow-Up was "a person who was a high school sophomore in spring 1980 but who was neither enrolled in high school nor a high school graduate or the equivalent at the time of the follow-up survey in spring 1982." Other school-leavers (such as early graduates) were separately surveyed.

More specifically, the HS&B First Follow-Up defined a dropout as a sample member who

> Has not attended school for the past month or more (not due to illness or accident), and does not intend to return.

Two dropout categories were provided as a further distinction within this general definition. One category encompassed dropouts neither enrolled in school nor attending a special program; the other embraced dropouts attending a special program (which might or might not be held within a school facility) such as a GED program or adult education courses. The reference period of "one month or more" is roughly equivalent to the twenty or more consecutive school days definition employed in the HS&B Base Year.

In HS&B, then, schools were asked to identify dropouts in 1980 to exclude them from the sample; and schools were asked to identify dropouts in 1982 so that these individuals could be pursued out-of-school and administered a special dropout questionnaire. Additional important school records data on completion is available through the high schools transcripts collected by HS&B. However, sample members were also asked about their school enrollment status on the HS&B questionnaires, so that there are multiple, and sometimes conflicting sources of

information. As Frase points out (1989, p.83), various dropout rates have been calculated for the HS&B sophomore cohort, owing to sample variation, differences in data sources, availability of alternative dropout definitions and decision rules about use of inconsistent sources, and so on.

3.1.2. NELS:88 Operational Definition of a Dropout.

In NELS:88, it is necessary to operationalize the concept of a dropout at two basic levels. First, in terms of an individual's status, who is a dropout and who is not?; second, in terms of instrument completion, who should be given the dropout questionnaire? Thus all sample members who were out of school in the spring of 1990 were to be administered the dropout questionnaire, but dropouts identified at earlier stages who had been back in school for some time were asked to complete the student questionnaire.

In the NELS:88 First Follow-Up, the following dropout definition was used:

1. an individual who, according to the school (if the sample member could not be located), or according to the school and home, is not attending school (= has not been in school for four consecutive weeks or more and is not absent due to accident or illness[2])

2. a student who has been in school less than two weeks after a period in which he or she was classified as a dropout (this individual would be classified as a stopout but administered the dropout, rather than the student, questionnaire; all other in-school stopouts would be administered the student questionnaire)

The NELS:88 definition differs somewhat from HS&B First Follow-Up's in that it is more purely behavioral in focus--it omits reference to the "intentions" of the sample member. (Interestingly, the HS&B Base Year definition did as well, offering instead the phrase "and is not expected to return"). It also differs somewhat from the proposed CCD definition. The CCD definition (see Kaufman and Frase, 1990, p. 27; and Hoffman 1990), if plans proceed unaltered, will be applied nationally in the 1991-92 school year, and will thus coincide with data collection for the NELS:88 Second Follow-Up. According to this definition:

> A school dropout is an individual who was enrolled in school at some time during the previous school year, was not enrolled at the beginning of the current school year, has not graduated from high school or completed an approved educational program, and does not meet any of the following exclusionary conditions:

---

[2] Analysts can use item 13 in the Student Questionnaire to look at chronic absentees. (This item reads: in the first half of the current school year, about how many days were you absent from school for any reason? Response options range from "None" to "21 or more").

* death;

* temporary absence due to suspension or illness;

* transfer to another public school district, private school, or a State- or District-approved education program.

For the purposes of this definition:

* A school year is the twelve-month period of time beginning with the normal opening of school in the fall

* An individual has graduated from high school or completed an approved education program upon receipt of formal recognition from school authorities

* A State or District approved education program may include special education programs, home-based instruction, and school-sponsored GED preparation

NELS:88 considers students enrolled in alternative programs as in school, as, generally speaking, do the participants in such programs[3]; the CCD definition

---

[3] In the NELS:88 First Follow-Up field test in the spring of 1989, we found that when students in alternative programs were asked to complete the dropout questionnaire, they found it difficult to answer some items because these questions implied that they had left school. Many of these students reported that they were not dropouts and still were in school, even though in most of these instances they could not have been referred to their alternative program unless they had dropped out of school. Our conclusion was that there may be some reluctance to identify oneself as a dropout when one is engaged in an alternative program toward school completion, and that the student questionnaire--if one is limited to but two questionnaires--may be the more appropriate survey instrument for alternative program participants to complete. Certainly the dropout population is highly differentiated. There are students who have left school, but there are also those who have returned to alternative or regular programs. In addition, there are students who are in alternative programs to prevent dropping out. Finally, there are significant numbers of students who are chronic truants. There are many gradations of disengagement along the continuum between in-school status and dropout status. Student self-reports of dropout status may be more or less reliable, depending on the definition of school-leaving that is being used, and the type and degree of disengagement from school exhibited by the respondent. There may also be some reluctance on the part of parents to label their children dropouts; it is unclear, for studies such as CPS, how often household respondents view GED and other alternative program participants as being "in school" and how often they do not. (Students in GED or alternative non-diploma instruction were administered the dropout questionnaire in HS&B.) We simply do not know whether early and late "alternative completers" tend to differ in their perceptions of their school enrollment status. For this reason, it is conceivable that different rules for assignment of in-school (student) and out-of-school (dropout) questionnaires may be appropriate at different stages of

differentiates between approved and unapproved programs. Because the district and state approved/disapproved/neither approved nor disapproved distinction is not necessarily meaningful to participants and members of their households, and because virtually indistinguishable programs may be approved in one state or locality and disapproved (or neither approved or disapproved) in another, employing this distinction for the national sample of NELS:88 did not seem a sensible direction to pursue. It might be contended that part of the purpose and benefit of a national study is provision of a classification scheme that overrides arbitrary or inconsistent state and local definitions; a student who would be classified as a dropout in one state and as in-school in another should be classified by a national study in a manner that enforces uniformity across state and district boundaries.

Nevertheless, it will be advisable in NELS:88 to specially flag all "alternative completers" so that members of this population can be specially studied, and comparability to broader dropout definitions maintained.

The October Current Population Survey asks the selected respondent about current enrollment status and enrollment one year previously; however, since a definition of being enrolled "in school" is not provided, households containing alternative completers may interpret the question as including or excluding alternative programs. CPS interviews a household member; hence the enrollment status report may be given by a proxy. It is not clear to what extent a parent is likely to offer a different response to such a question than would a school noncompleter or alternative completer. Dropping out of school is a socially undesirable behavior, and social desirability norms may influence survey responses. In many cases where social desirability bias is feared, one obtains more accurate reports from a secondary source or proxy than from the primary source respondent. However, the limited evidence available suggests that any knowledge deficit or social desirability bias in reporting dropout status may affect parents/household members more than out-of-school youth. For example, Mohadjer, Brick and West (1990), reporting on the correspondence between the responses of knowledgeable household respondents and youths in the National Household Education Survey field test, derived smaller dropout estimates from household interviews than from youth interviews. Differences were especially pronounced for event dropout rates, less so for status dropout rates, and for younger dropouts (presumably under-reported by household proxies) as contrasted to older. There is also the possibility that proxy reports differ in their reliability by racial/ethnic group.

Another problematic feature of the CPS is that it permits no "missing" data--thus, for example, interviewers are instructed to not accept a response such as "Don't Know" but to probe until an answer matching the response categories is achieved. The effect of this practice on the accuracy of estimates is unknown, but the methodological literature generally points to difficulties with such an approach. Most studies therefore employ reserve codes to accommodate item refusal, don't know responses, and other sources of missing data at the item

---

a longitudinal study. The NELS:88 Second Follow-Up can be made consistent in this respect with its prior round, or with the twelfth grade round of HS&B, but not with both.

level, hoping to learn thereby more about the reasons for nonresponse and to minimize incentives for interviewers to infer or manufacture responses that are not readily forthcoming from the respondent.

HS&B, in contrast both to CCD and NELS:88, considered students enrolled in alternative programs with a GED focus--whether approved or unapproved--as a special category of dropout. Since, however, it separately distinguished such students, it permits the recombination of data categories to fit multiple definitions. The general strategy of recognizing multiple distinctions, of providing various sorting and classifications categories--as opposed to forcing categorization across a dropout/not a dropout dichotomy--gives analysts greater flexibility and maximizes comparability to alternative data sources. Such a strategy also no doubt better reflects the intermediate stages in the continuum of statuses and dispositions between full engagement with and disengagement from schooling, as experienced by students in regular schools, alternative programs, and school non-completers. This strategy--preserving classification categories that are important to cross-study comparability, and that specifically address the school engagement/disengagement experience of students--should be vigorously pursued in NELS:88 and other national studies concerned with production of dropout statistics.

Frase points out (1989, p.80) that "the dropout definitions embedded in the existing data sources -- CPS and HS&B -- are neither consistent with one another nor with the new definition that NCES is trying to develop." As we have seen, the NELS:88 operational definition of a dropout differs in some respects also from the CCD definition and even from the HS&B definition.

Since consistency and comparability between definitional approaches is in the main desirable, Frase's conclusion is unsettling. If some of our most important data sources are definitionally inconsistent with each other, and if new studies such as NELS:88 should ideally produce data comparable to historical as well as consciously redefined sources, some difficulty may be anticipated in simultaneously doing justice to the requirements of a longitudinal study in the 1990s, whilst providing cross-cohort comparability to a major longitudinal study of the 1980s, as well as comparison points to non-longitudinal data sources (such as recurrent household surveys and administrative records systems). Perhaps the most useful response to this quandary would be to attempt to provide a sufficiently rich (and "rich" in one of its prominent meanings may be identified with tolerance of ambiguity and contradiction) information source that various definitions might be constructed and reconstructed to match to a variety of external sources. Such an approach surely would maximize the utility of a database such as NELS:88.

3.2: The Method of double confirmation of dropout status in NELS:88.

Both under and over-reporting are significant problems in administrative records, quite apart from problems in the variation of records definitions from site to site. Sometimes a student may be presumed to be a dropout but actually be a transfer. As LeCompte and Goebel point out (1987, p. 254), some states require that schools obtain academic transcripts for incoming transfer students, but others do not. Such requirements, moreover, are otiose if a student has

transferred from the public to the private school sector, across state lines, or out of the country. In other cases schools may be reluctant to classify an out-of-school student as a dropout, whether because the individual is below legal school-leaving age or for other reasons, or there may be reporting inaccuracies which lead to misclassifications.

The methodological approach of NELS:88 is to systematically confirm with the household dropout status as reported by the school. Thus, if a school reports that a given student is a dropout, the household will be asked to verify this status. If the parent reports that the student has transferred to another school, then the in-school status of the student can be verified from the records of the newly-named source. Of course, in school records are not wholly definitive guides to the enrollment status of individuals, neither are household reports. A parent may mistakenly believe that a student is attending school, or may misreport on this sensitive topic owing to embarrassment at the child's behavior.

Since much error resides in administrative records data upon which school personnel rely and in the subjective perceptions of household members as well, a major problem of the dual source approach is how to resolve contradictory enrollment status reports. Conflict between household and school sources reflects various degrees of self-interest and subjectivity in reporting, but also complex questions of interpretation and meaning, surrounding such issues as what it means to be "in school" given the availability of alternative programs and the complex continuum of statuses and behaviors that reside between full engagement with and disengagement from school.

In order to elicit consistent information from school and household data sources and in order to code and classify information provided by these sources, it was necessary to further define what it means to be in school. To this end, anyone who was not enrolled in a "traditional" school but was taking academic courses or receiving academic instruction (including those who had left school for home-based instruction or were receiving academic instruction while incarcerated) was considered a student. On the other hand, those below legal school-leaving age who met the behavioral criteria for dropout status were considered dropouts.

Despite the difficulty of contradictory reports that arises whenever multiple sources are consulted, there can be little doubt but that using multiple sources to "correct" each other should lead to superior estimates in the end, if sensible decision rules are consistently used to resolve such conflicts. However, there is a further difficulty here--namely, that beginning with the school's records and using the household to verify dropout status is a good way to limit school over-reporting of dropout status (which is a frequent occurrence, normally because of missed transfer-out status). But, unless one also checked with the household about those individuals identified as in-school students, one has no mechanism to correct underreporting of dropout status by the school. To a large degree this problem is circumvented by not relying on the school's definition of dropping out--one instead grounds judgments in the attendance record. But there is variation as well in what attendance means from place to

place, and in the quality of attendance records.

In addition, specific decision rules are required to adjudicate between conflicting claims of household and school. In the NELS:88 First Follow-Up, if the school said the sample member was a dropout and the home said the sample member was a student, field staff informed the household member/respondent that the school had indicated that the sample member had left school as of 00/00/00, and then asked, did the student transfer to another school, if so which one and when. If the supposed destination school after transfer denied that the sample member was enrolled, the household report was discounted. If the household maintained that the student was regularly attending the original school, and school records indicated otherwise, the household report was discounted. The animating assumptions here were that the social undesirability of dropping out might make it difficult for a parent to admit that their child was a dropout, and that a parent might be in ignorance of their offspring's school-attendance behavior.

4. Design Issues in the Longitudinal Approach to Dropout Statistics.

4.1 Limitations in the HS&B Approach

One weakness in the High School and Beyond design is that it began with second semester tenth graders; many students drop out before reaching the spring term of tenth grade.

A second limitation of the HS&B design is that it excluded certain categories of students: those who dropped out in the course of tenth grade, those with language barriers to participation or with physical or mental barriers to participation. These excluded students do not enter into the cohort dropout rate obtained from HS&B.

A third weakness of the HS&B design is that it does not provide definitive enrollment status information for the full sample. Analysts have data for those who completed a student questionnaire, but do not have enrollment data for nonparticipants. Participation rates in the HS&B First Follow-Up were extraordinarily high--96 percent. Nevertheless, there may have been "hidden dropouts" in the population of students (as defined by the school) who did not participate despite survey days and repeated make-up days.

4.2 Dropout Statistics in NELS:88: Remedying HS&B's Weaknesses

4.2.1. An Earlier Beginning Point. NELS:88 began with eighth graders, thus providing a baseline immediately prior to entry into secondary school. Even this earlier beginning point may not be wholly satisfactory. Insofar as a longitudinal studies approach fails to begin early enough (as was certainly the case with HS&B), its dropout estimates must be taken as underestimates and one must turn to household studies and administrative records to supplement longitudinal data. NELS:88, in starting with eighth graders, largely, but not entirely, corrects this deficiency in HS&B. Frase (1989, p.22), using Bureau

of the Census CPS data, shows 12 percent of dropouts have "completed six years of elementary school at most"--presumably, this portion of the dropout population would be missed by a study such as NELS:88, with its beginnings in an eighth grade cohort. As Frase also points out, over 31 percent of Hispanic dropouts age 16-24 had completed six or fewer years of schooling. The biasing effect on NELS:88 estimates of pre-eighth grade school-leaving may, then, be decidedly more pronounced for some subgroups than for others.

4.2.2 **Remedying the Effects of Ineligibility Exclusions.** A purely records-based approach does not have to contend with undercoverage problems stemming from survey exclusion, nor do household surveys, though depending on the methodology of the survey or the universality and accuracy of the records system, undercoverage problems may plague these approaches too, in somewhat different guise. Be that as it may, an enduring and assuredly non-trivial problem in national studies, whether longitudinal sample surveys such as HS&B and NELS:88, or repeated cross-sectional assessments such as NAEP, is that certain kinds of schools and categories of students are excluded, leading to possible undercoverage bias. If one has reason to think that students in these excluded categories leave school at either a higher or lower rate than students in general, then ineligibility factors are a potentially quite significant problem for national dropout estimation. Subgroup estimates may be particularly affected. For example, exclusion owing to language barrier may particularly affect groups with high recent immigration rates to the United States, including many Hispanic and Asian subgroups. Additionally, excluded members of the subgroup are likely to differ in important respects from included members. Thus, for example, Bean and Tienda (1987) report that it is precisely those Hispanics with lesser English language proficiency who are least likely to complete high school.

Certain kinds of specialized schools were systematically excluded from the NELS:88 sample such as Bureau of Indian Affairs schools and special schools for the handicapped. Also, children educated at home or in other non-school settings were excluded. The effect of these exclusions on NELS:88 estimates is thought to be very slight (see Spencer et al., 1990, p.3).

More important to dropout estimation was the exclusion of certain categories of students. NELS:88 maintained eligibility criteria essentially similar to those of HS&B (see Appendices B and C for ineligibility criteria in HS&B and NELS:88). As a result, approximately 5.34 percent of the potential sample was excluded for reasons of language barriers, or physical or mental disabilities. (A good comparison point is 1988 NAEP at eighth grade level; its 5.3 percent student sample exclusion rate is remarkably similar to the NELS:88 rate).

However, in order to produce a correction factor for NELS:88 eighth grade cohort national dropout estimates, a subsample of the Base Year ineligible students is being pursued in the NELS:88 First Follow-Up. School enrollment status information and additional demographic data will be obtained for this group. Also, their eligibility status will be re-examined, and those whose status has changed (for example, students excluded in 1988 for lack of sufficient proficiency in English, who in the meantime have become sufficiently proficient

to complete the survey instruments) will be added to the First Follow-Up sample, along with freshened students. (For a full discussion of the followback study of excluded eighth graders, see Ingels, 1991).

4.2.3. Screening Ambiguous Pendings; Calculating a Cohort Dropout Rate Based on an Expanded Sample Definition

The NELS:88 First Follow-Up also screened for enrollment status a fifty percent subsample of all pending dispositions in which dropouts may be hidden (specifically, nonrespondents--sample members not identified as dropouts by their schools but who did not participate at either the initial survey session or at subsequent makeup days; students who were not located at the expected school in the data collection phase, and required further locating). The rationale for screening nonrespondents is that later information from records sources may frequently supersede the initial Phase 3 categorizations given to sample members by schools. (That is, there may be a gap between the time a student leaves a school, and the time when the origin school receives a request for academic transcripts from the destination school; in the meantime, the former student's status is unknown, and he/she may mistakenly be assumed to be a dropout.) There is therefore some benefit in revisiting the question of enrollment status at a later date when the whereabouts and status of missing students/dropouts may more accurately be ascertained. Unknown dropouts as well as transfers to other schools may of course be concealed in the population of sample members whose whereabouts and status were unknown. All ambiguous pendings should therefore be subject to further investigation, following the double confirmation (school and household) process. In this way, definitive enrollment status information is obtained for nonparticipants as well as for survey participants. By virtue of this component and the followback of base year ineligibles, it is possible to compute an expanded sample weight encompassing the entire eighth grade cohort and to calculate its enrollment status two years later.

In general, the approach of HS&B--to ground estimation in sample members who have completed the student questionnaire--is being supplemented in NELS:88 by estimation that draws on school enrollment status and demographic data for nonparticipants and excluded students as well, thus facilitating more accurate national estimates of a cohort dropout rate.

5. Conclusions

It is extremely important to be able to compare dropout statistics across surveys and records systems, and to ascertain the reasons for differences in estimates produced by each. More uniform dropout definitions will enhance the overall utility of administrative records; school-based longitudinal studies will benefit substantially from this greater precision and stability of meaning, insofar as they rely on schools and other records sources for part of the data that they collect. At the same time, a multiplicity of approaches with a common definitional referent will be most likely to produce accurate and useful

estimates. Studies such as NELS:88 can underwrite both causal modeling of dropout phenomena, and improved descriptive statistical reporting. Studies such as NELS:88 can enrich trend analyses by collecting data elements necessary to historical comparisons, and can enrich cross-sectional statistics by maximizing comparability to more standardized records definitions of school leaving. The simultaneous movement toward more uniform definitions, and toward enhanced powers of estimation in longitudinal education studies, give substantial promise of an improved statistical system for monitoring the critical phenomenon of school leaving and school completion.

REFERENCES (for text and appendices)

Barro, Stephen M., and Kolstad, Andrew. 1987. Who Drops Out of High School?: Findings from High School and Beyond. Washington, D.C.: National Center for Education Statistics.

Bean, Frank D., and Tienda, Marta. 1987. The Hispanic Population of the United States. New York: Sage.

Frase, Mary J. Dropout Rates in the United States: 1988. 1989. Washington, D.C.: NCES. (Analysis Report 89-609).

Hafner, Anne L.; Ingels, Steven J.; Schneider, Barbara; and Stevenson, David L.. 1990. A Profile of the American Eighth Grader. Washington, D.C.: National Center for Education Statistics. (NCES 90-458).

Hafner, Anne L. 1990. Dropout Statistics in NELS:88. Boston, MA: Annual meeting of the American Educational Research Association.

Hoffman, Lee McGraw. 1990. Issues in Developing Comparable National Dropout Statistics Through the Common Core of Data Survey. Boston, MA: Annual meeting of the American Educational Research Association.

Ingels, Steven J.; Abraham, Sameer Y.; Frankel, Martin R.; Karr, Rosemary; Rasinski, Kenneth A.; Spencer, Bruce D. 1990. NELS:88 Base Year Student Survey User's Manual. Washington, D.C.: National Center for Education Statistics. (NCES 90-464).

Ingels, Steven J. 1991. The Problem of Excluded Baseline Students in a School-Based Longitudinal Study: Correcting National Dropout Estimates and Accommodating Eligibility Change Over Time. Chicago: Annual meeting of the American Educational Research Association, Session 34.15.

Kaufman, Phillip, and Frase, Mary J. 1990. Dropout Rates in the United States: 1989. Washington, D.C.: National Center for Education Statistics. (NCES 90-659).

Kish, Leslie. Survey Sampling. 1965. New York: John Wiley.

Kominski, Robert. 1990. Estimating the High School Dropout Rate. Demography, 27(2).

LeCompte, Margaret D., and Goebel, Stephen D. 1987. "Can Bad Data Produce Good Program Planning? An Analysis of Record-Keeping on School Dropouts." Education and Urban Society, 19(3), 250-268.

Mohadjer, Leyla; Brick, Mike; and West, Jerry. 1990. "Proxy Respondents and Measurement Errors for Statistics on Dropouts." International Conference on Measurement Error in Surveys, Tucson, Arizona, November 1990.

Spencer, Bruce D.; Frankel, Martin R.; Ingels, Steven J.; Rasinski, Kenneth A.; and Tourangeau, Roger. 1990. NELS:88 Base Year Sample Design Report. Washington, D.C.: National Center for Education Statistics.

Yates, F., 1948, "Systematic Sampling," Philosophical Transactions of the Royal Society (A; 241, 345-377).

Appendix A: Sample Freshening.

Sample freshening procedures are employed in order to give some chance of selection into the NELS:88 First Follow-Up (1990) and Second Follow-Up (1992) to "out of sequence" students, that is, to 1990 tenth graders and 1992 twelfth graders who were not in the eighth grade in the 1987-88 school year. The sample freshening procedure was devised by Martin R. Frankel as an application of the "half-open interval procedure" for minimizing undercoverage. The half-open interval technique has previously been applied primarily in the context of household surveys to correct for missed dwellings in an area frame (see Yates, 1948; and Kish, 1965). However, the procedure can be adapted to encompass "missed" (that is, not part of the eighth grade cohort because not in eighth grade two years before) tenth grade student populations, so that all portions of the population of tenth graders in 1990 (and twelfth graders in 1992) can be given a chance of selection into the NELS:88 sample. Operationalization of the freshening procedure for the First Follow-Up is as follows:

1. In the fall of 1989, NORC Survey Representatives obtained the complete tenth grade student roster from high schools selected for inclusion in the sample, in order:

    a) to ascertain that students expected to be enrolled are in fact enrolled; and
    b) to perform the student linking operations on each school's roster that form the basis for calculating the selection probabilities for "freshened" cases drawn from the tenth grade rosters.

2. Tenth grade rosters were examined by the Survey Representatives within the school building and each NELS:88 eighth grade selection present on the roster was linked down to the next student (referred to as a "linked" student) on the list.

3. The linked student was "screened" to determine whether he or she was enrolled in the eighth grade in the United States two years earlier. A review of the names of linked students by a school administrator able to consult students' cumulative files was sufficient to make this determination.

4. Students who were eighth graders in this country two years prior are ineligible for the freshened sample. The linking process stops with that ineligible student.

5. Each linked student who was not in the eighth grade in this country two years earlier was selected into the "freshened" sample component. If a linked student was selected, the immediately following student on the list was also considered "linked" to the NELS:88 sample member.

6. Steps 3, 4, and 5 were repeated for each linked student, with the process continuing until a student linked to each NELS:88 eighth grader was determined to be ineligible.

Appendix B: HS&B BASELINE SOPHOMORE COHORT INELIGIBILITY CRITERIA

Within the eligible schools, a sophomore was defined as a student who expected to complete his/her tenth grade course work between April 1, 1980 and August 31, 1980. This was to include those students who might be held back or who might repeat tenth grade, but to exclude students dropping out before administration of the HS&B questionnaire in the spring of 1980.

NORC asked each selected school to provide a list of its tenth grade students, as defined above. All students defined as being eligible for sample selection were included in the sampling frame; conversely, all students defined as ineligible for sample selection were removed from the frame.

Additionally, however, a student was considered ineligible when that student:

- Was a foreign exchange student
- Transferred out of the selected school. (A transfer student was defined as a student who had left the school and whose records were requested for a new school).*
- Died.
- Would be unavailable until after August 31, 1980.
- Was listed on the roster in error.
- Had become a drop-out or lost student since he was selected. Such a student would have to have been out of school for 20 or more consecutive days and was not expected to return.**
- Was physically or mentally unable to participate in the survey. ***

*Transfers-out were not directly replaced. (However, as in NELS:88, all transfers-in were given a chance of selection into the sample). All other categories above in HS&B base year led to replacement by other students from the roster. (HS&B substituted students for the following cases: dropout, listed in error, language barrier, too ill [mentally, physically], in jail, unavailable entire field period, expelled, and deceased.) No substitution of students was done in the NELS:88 base year.

**Dropouts meet the 20 consecutive days criterion, are at least 16 years of age, and are not expected to return to school. Lost students are dropouts in all respects except that they are not 16 years of age.

***While this category was used to cover linguistic exclusion also, a Spanish language version of the questionnaire was provided so that students whose primary language competence was in Spanish would not be excluded. However, only 36 sophomores and 8 1980 HS&B seniors elected to complete the instrument in Spanish. Unfortunately, some native speakers of Spanish

are literate in neither Spanish nor English. Also, some students with limited proficiency in English and greater proficiency in Spanish may nonetheless prefer not to call attention to their LEP status.

Appendix C: NELS:88 Base Year Exclusion Criteria.

Seven ineligibility codes defining categories of excluded students were employed at the time of student sample selection:

A - attended sampled school only on a part-time basis, primary enrollment at another school.

B - physical disability precluded student from filling out questionnaires and taking tests.

C - mental disability precluded student from filling out questionnaires and taking tests.

D - dropout: absent or truant for 20 consecutive days, and was not expected to return to school.*

E - did not have English as the mother tongue AND had insufficient command of English to complete the NELS:88 questionnaires and tests.

F - transferred out of the school since roster was compiled.

G - was deceased.

*However, the small number of students who dropped out in the Base Year between drawing of the sample (between October of 1987 and January of 1988) and survey days (between February and June of 1988) were specially followed (see Ingels et al. 1990, App. E.).